CLAIMS

What is claimed is:

1. A method of classifying at least one test time series data set utilizing at least one training time series data set comprising the steps of:

5   creating wavelet decomposed segments of the at least one training time series data set;

discretizing the wavelet decomposed segments of the at least one training time series data set;

determining one or more frequent patterns in the discretized wavelet decomposed

10   segments of the at least one training time series data set;

determining rules from the one or more frequent patterns in the discretized wavelet decomposed segments; and

classifying one or more frequent patterns in the at least one test time series data set according to the rules.

15   2. The method of claim 1, wherein the step of determining one or more frequent patterns is based on user-specified support and width values.

3. The method of claim 1, wherein the step of determining one or more frequent patterns is an iterative process using frequent k-patterns to generate candidate (k+1)-patterns.

4. The method of claim 1, wherein the step of discretizing the wavelet decomposed segments comprises creating a set of categorical records on which the one or more frequent patterns may be found.

5. The method of claim 1, further comprising the step of determining rules of a user-specified level of confidence.

6. The method of claim 5, further comprising the step of sorting the rules in order of decreasing confidence.

7. The method of claim 5, further comprising the step of determining a default class using the at least one training time series data set and the rules generated.

8. The method of claim 7, wherein the default class is a majority class of data instances which do not match any rule.

9. The method of claim 1, wherein the step of classifying the one or more frequent patterns comprises the steps of:

performing wavelet decomposition to a test instance of the at least one test time series data set;

discretizing the wavelet decomposed segments of the test instance; and

finding a rule having a pattern that is a subpattern in the discretized wavelet decomposed segments of the test instance.

10. The method of claim 6, wherein the step of classifying the one or more frequent patterns comprises reporting a class label of a highest confidence rule matched by a given data record of the at least one test time series data set.

11. The method of claim 7, further comprising the step of reporting the default class as a class label when there is no matching rule for a data record of the at least one test time series data set.

12. Apparatus for classifying a time series data, the apparatus comprising:

a memory; and

at least one processor, coupled to the memory operative to: (i) create wavelet decomposed segments of the training time series data set; (ii) discretize wavelet decomposed segments of the training time series data set; (iii) determine one or more frequent patterns in discretized wavelet decomposed segments of the training time series data set; (iv) determine rules from one or more frequent patterns in the discretized wavelet decomposed segments; and (v) classify one or more frequent patterns in the test time series data set according to the rules.

13. The apparatus of claim 12, wherein the operation of determining one or more frequent patterns is based on user-specified support and width values.

14. The apparatus of claim 12, wherein the operation of determining one or more frequent patterns is an iterative process using frequent k-patterns to generate candidate (k+1)-patterns.

15. The apparatus of claim 12, wherein the operation of discretizing the wavelet decomposed segments comprises creating a set of categorical records on which the one or more frequent patterns may be found.

16. The apparatus of claim 12, wherein the at least one processor is further operative to determine rules of a user-specified level of confidence.

17. The apparatus of claim 16, wherein the at least one processor is further operative to sort the rules in order of decreasing confidence.

18. The apparatus of claim 16, wherein the at least one processor is further operative to determine a default class using the at least one training time series data set and the rules generated.

19. The apparatus of claim 18, wherein the default class is a majority class of data instances which do not match any rule.

20. The apparatus of claim 12, wherein the operation of classifying the one or more frequent patterns comprises:

performing wavelet decomposition to a test instance of the at least one test time series data set;

discretizing the wavelet decomposed segments of the test instance; and

finding a rule having a pattern that is a subpattern in the discretized wavelet decomposed segments of the test instance.

21. The apparatus of claim 17, wherein the operation of classifying the one or more frequent patterns comprises reporting a class label of a highest confidence rule matched by a given data record of the at least one test time series data set.

22. The apparatus of claim 18, wherein the at least one processor is further operative to report the default class as a class label when there is no matching rule for a data record of the at least one test time series data set.

23. An article of manufacture for classifying at least one time series data set, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

creating wavelet decomposed segments of the at least one training time series data set;

discretizing the wavelet decomposed segments of the at least one training time series data set;

determining one or more frequent patterns in the discretized wavelet decomposed segments of the at least one training time series data set;

determining rules from the one or more frequent patterns in the discretized wavelet decomposed segments; and

classifying one or more frequent patterns in the test time series data set according to the rules.

24. The article of manufacture of claim 23, wherein in the step of determining one or more frequent patterns is based on user-specified support and width values.